

DOI:10.3969/j.issn.1000-1298.2010.08.038

基于改进型 FCM 算法的牛肉大理石花纹提取方法*

仇金宏¹ 沈明霞¹ 彭增起² 梁琨¹ 吴海娟¹ 史杰²

(1. 南京农业大学工学院, 南京 210031; 2. 南京农业大学食品学院, 南京 210095)

【摘要】 提出了一种基于改进型模糊 C 均值聚类算法的牛肉大理石花纹提取方法。该方法结合了快速模糊 C 均值(FCM)聚类算法,对传统 FCM 算法中的隶属函数、聚类数 C 和初始聚类中心点选取方法进行了优化。试验表明,该方法使牛肉大理石花纹提取的准确度由 76.2% 提高到 85.7%。

关键词: 牛肉 大理石花纹 提取 模糊 C 均值

中图分类号: TS251.5⁺2; TP391.41 文献标识码: A 文章编号: 1000-1298(2010)08-0184-05

Beef Marbling Extraction Based on Modified Fuzzy C-means Clustering Algorithm

Qiu Jinhong¹ Shen Mingxia¹ Peng Zengqi² Liang Kun¹ Wu Haijuan¹ Shi Jie²

(1. College of Engineering, Nanjing Agricultural University, Nanjing 210031, China)

(2. College of Food Engineering, Nanjing Agricultural University, Nanjing 210095, China)

Abstract

Beef marbling extraction method based on modified fuzzy C-means clustering algorithm was proposed. The proposed method combined with the fast fuzzy C-means clustering algorithm. The membership function, number of clusters C and the selection of initial cluster center in traditional fuzzy C-means clustering algorithm were modified. Experiments showed that, with the proposed method, the accuracy of marbling extraction was increased from 76.2% to 85.7%.

Key words Beef, Marbling, Extraction, Fuzzy C-means

引言

在采用机器视觉技术进行牛肉品质的自动评定时,一般主要依据牛肉的大理石花纹^[1-3],所以牛胴体横切面中大理石花纹提取的准确程度对后续大理石花纹等级的自动评定有着重要的影响^[4-6]。牛胴体横切面图像具有模糊性和不确定性,其中包括灰度和几何的模糊性以及不确定知识,从而使大理石花纹提取的准确程度无法得以保证,所以模糊聚类方法是一种有效的牛肉大理石花纹提取方法。目前应用较广泛的是模糊 C 均值(FCM)聚类算法^[4,7-8]。

然而,FCM 算法在使用时需引入一定的归一化条件,因此该聚类方法在牛胴体横切面图像样本不理想的情况下可能会导致不理想的聚类结果。而且,像大多数非线性优化问题一样,FCM 算法对数据聚类效果的优劣,在很大程度上依赖于参数初值,如聚类数 C 、初始聚类中心点、指数权值 m 等的选择,也正是这些因素制约着 FCM 算法的进一步应用。Chiu Spheh L 部分解决了聚类数 C 和初始聚类中心点选取的问题,提出了一种自适应决定聚类数 C 以及初始聚类中心的方法^[9-10]。但是它在解决 FCM 的初值问题时,又引入了一些对聚类效果有影响的新问题,致使 FCM 算法不适合直接用于牛肉大

收稿日期: 2010-03-16 修回日期: 2010-04-17

* 农业部肉牛产业技术体系资助项目(080600231,080600232)

作者简介: 仇金宏,硕士生,主要从事机器视觉与图像处理在农产品品质检测领域的应用研究,E-mail: njauqjh@gmail.com

通讯作者: 沈明霞,教授,博士生导师,主要从事机器视觉和信息农业研究,E-mail: mingxia@njau.edu.cn

理石花纹的精确提取。为此,本文针对牛胴体横切面图像的特点,提出用改进型模糊 C 均值(FCM)聚类算法——FCM-SQP 算法提取牛肉的大理石花纹。

1 灰度和空间特征的快速 FCM 算法

1.1 FCM 算法

图像 FCM 聚类分割实现方法是根据图像像素和聚类中心的加权相似性测度,对目标函数进行迭代优化以确定最佳聚类。FCM 聚类是通过最小化隶属度矩阵 U 和聚类中心 V 的目标函数 $J_m(u, v)$ 来实现的

$$J_m(U, V; X) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^m \|x_k - v_i\|_A^2 \quad (1)$$

$(i=1, 2, \dots, c; k=1, 2, \dots, n; 2 \leq c \leq n)$

其中 $U = \{u_{ik}\}$ $V = \{v_1, v_2, \dots, v_c\}$
 式中 U ——满足上述约束条件的 $c \times n$ 隶属度矩阵
 u_{ik} ——第 k 个灰度值 x_k 对应的第 i 个隶属度值
 V —— c 个聚类中心点集

由于在提取牛肉大理石花纹的过程中需把牛胴体横切面有效区域图像按灰度值分成脂肪 v_1 和肌肉 v_2 两类,故在牛肉大理石花纹提取的过程中选择 $c=2; m \in [1, +\infty]$ 为模糊加权指数,当 $m=1$ 时,模糊聚类就退化为硬 C 均值聚类。Nikhil 等人的研究表明, m 的最佳选择范围为 $[1.5, 2.5]$,通常 $m=2$ 是比较理想的取值^[11]。 x_1, x_2, \dots, x_n 是牛胴体横切面有效区域图像 n 个像素点的灰度值, n 即是牛胴体横切面有效区域图像的像素点总数; $\|x_k - v_i\|_A^2$ 又可表示为 $|x_k - v_i|^2$,表示 x_k 与 v_i 之间的归一化距离,即 k 个样本到第 i 类中心的距离。

1.2 快速 FCM 算法

在进行牛肉大理石花纹提取的过程中,由于图像样本数据量 T 很大(一幅 640×480 的图像, $T=307\,200$),故 FCM 算法的收敛速度很慢。李明^[11]等对提高 FCM 算法速度进行了研究,利用将图像从像素空间映射到一维灰度直方图特征空间的方法,提出了一种快速 FCM 算法。在该快速 FCM 算法中,新的快速 FCM 目标函数和聚类中心为

$$u_{is} = \frac{1}{\sum_{j=1}^c \left(\frac{\|s - v_i\|_A}{\|s - v_j\|_A} \right)^{2/(m-1)}} \quad (2)$$

$$v_i = \frac{\sum_{s=1}^{S-1} (u_{is})^m H(s) S}{\sum_{s=1}^n (u_{is})^m H(s)} \quad (3)$$

根据文献[11]的研究可以类推得知:利用该快

速 FCM 算法进行牛肉大理石花纹的提取,其结果与利用 FCM 算法进行牛肉大理石花纹提取的结果完全一样。同时,由于改进的快速 FCM 算法中引入了直方图统计特性,使得算法处理的数据个数大大减少。若图像的大小为 640×480 ,灰度级 $S=256$ 。用此算法提高的效率为 $640 \times 480 / 256 = 1\,200$ 倍。由此可知:利用该快速 FCM 算法可以明显地提高牛肉大理石花纹的提取速度。

2 改进型模糊 C 均值(FCM-SQP)聚类算法

2.1 FCM 均值聚类算法中隶属度函数的改进

由于引入归一化条件: $\sum_{i=1}^c u_{ik} = 1 (1 \leq k \leq n)$,

当某个样本远离各类的聚类中心点时,本来它严格属于各类的隶属度都很小,但是由于归一化条件的要求,将会使它对各类都有较大的隶属度,从而影响了迭代的最终结果。因此,上述归一化条件下 FCM 聚类方法在样本集不理想的情况下可能导致不理想的聚类结果。为此,提出可选放松度的归一化条件,使所有样本对各类的隶属度总和为 $\sum_{i=1}^c \sum_{k=1}^n (u_{ik}) = n\ell$ (n 即为样本总数, $\ell \in (1/n, 1]$, ℓ 可根据相应要求进行选择),即

$$\sum_{i=1}^c \sum_{k=1}^n (u_{ik}) = n\ell \quad (4)$$

在此新的约束条件下,聚类中心的计算式 v_i (式(3))保持不变,而隶属度函数 u_{ik} 的计算式(式(2))则变为

$$u_{is} = \frac{n\ell}{\sum_{j=1}^c \left(\frac{\|s - v_i\|_A}{\|s - v_j\|_A} \right)^{2/(m-1)}} \quad (5)$$

2.2 FCM 均值聚类算法中聚类数 C 和初始聚类中心点选取方法的改进

文献[9]中提出的一种自适应决定聚类数 C 以及初始聚类中心的方法,它在解决 FCM 初值问题应用中效果并不十分显著的同时还引入了对聚类效果有影响的 2 个新问题:聚类中心产生振荡现象和算法应用过程中产生数据假聚类现象。针对聚类中心产生振荡现象的问题,文献[10]的解决办法是把已经成为聚类中心的点 x_{c-1} 从形成下一个聚类中心 x_c 的候选数据点中删除掉。但是,它在控制聚类中心振荡现象的发生时,无法在形成下一个聚类中心的候选数据点中剔除那些数值上无限接近于 x_{c-1} 的伪聚类中心点。针对算法应用过程中产生数据假聚类现象的问题,文献[10]的解决办法是取到所有已经生长的聚类中心点的距离总和为最大值 l_{\max} 的数据

点,用具有最高的聚类候选值的标准来进行聚类,但是,那些到所有已经生长的聚类中心点的距离总和无限接近于 l_{\max} 的点仍存在着十分明显的影响。为此本文在借鉴文献[10]的基础上,给出了解决上述问题的相应改进方法。

(1) 对聚类中心振荡现象的控制

在实际运用中,文献[9]中提出的算法很容易出现2个聚类中心 x_1, x_2 交替连续的情况,亦即形成聚类点为 $x_1, x_2, x_1, x_2 \dots$ 。解决的办法是把已经成为聚类中心的点 x_{c-1} 和数值上无限接近于 x_{c-1} 的伪聚类中心点从形成下一个聚类中心 x_c 的候选数据点中删除掉。

(2) 对数据假聚类现象的克服

文献[9]中提出判断下一个聚类中心点的标准是距离第 i 个聚类中心 x_i 最大的数据点具有最高的聚类候选值。这样就容易在形成聚类时产生聚类中心点不符合聚类要求的假聚类现象,亦即对 x_1 来说, x_2 是符合要求的,由 x_1 产生的聚类中心距离 x_1 是最远的。但是按照原算法,在由 x_2 产生下一个聚类中心 x_3 时, x_3 就很可能出现距离 x_2 很远,但是同时距离 x_1 却很近的现象。显然这种在 x_1 点附近形成的新聚类中心,却从整体上是符合数据聚类要求的。为此,采取到所有已经生长的聚类中心点距离总和为最大值 l_{\max} 和无限接近于 l_{\max} 的数据点,用具有最高的聚类候选值的标准进行聚类。

2.3 改进型模糊C均值聚类(FCM-SQP)算法实现

为实现改进型模糊C均值聚类算法能适用于牛肉大理石花纹的提取,借鉴文献[10]和文献[11]中的相关算法,并对其中存在的不足进行了改善,构建了改进型模糊C均值聚类(FCM-SQP)算法如下:

(1) 选取 c, m, ε ,根据样本总数 n 和实际情况选取式(5)中的 ϕ ,初始迭代时, $b=0$,以每一个数据点作为一个潜在的聚类中心点。

(2) 利用快速FCM算法对牛肉图像进行分割。

(3) 对每个数据点 x_i 定义测度

$$Q_i = \sum_{j=1}^n e^{-\phi \|x_i - x_j\|^2} \quad (6)$$

其中

$$\phi = \frac{4}{r_d^{2/(m-1)}}$$

式中 r_d ——正常数

Q_i —— x_i 到其他任一个数据点距离的函数常数

计算各 Q_i 值,选择对应 Q_i 值最高的点 x_i 作为第一个聚类中心 x_1^* 。

(4) 根据式(5)计算像素点的隶属度 u_{bs} 。

(5) 把已经成为聚类中心的点 x_{k-1} 和数值上无

限接近于 x_{k-1} 的伪聚类中心点从形成下一个聚类中心 x_k 的候选数据点中删除掉。

(6) 计算下一个聚类中心点

$$Q_{i+1} = Q_i - Q_i^* e^{-\varphi \|x_i - x_i^*\|^2} \quad (7)$$

其中 $\varphi = \frac{4}{r_e^2}$ $r_e = 1.5r_d$

距 x_i^* 越近的数据点,其相应的 Q_i 就越小,故而成为下一个聚类中心点的可能性就越小。

(7) 取到所有已经生长的聚类中心点距离总和为最大值 l_{\max} 和无限接近于 l_{\max} 的数据点,用具有最高的聚类候选值标准来进行聚类。进行迭代运算,计算第3、4、...个聚类中心点,经过 t 次聚类后, $Q_{i+1} = Q_i - Q_i^* e^{-\varphi \|x_i - x_i^*\|^2}$,直到 $Q_i^* < \varepsilon Q_1^*$ ($\|Q_{b+1} - Q_b\| < \varepsilon$)为止;否则转向步骤(3), $b = b + 1$ 。最终得到收敛后的隶属度和聚类中心。

3 算法的试验结果与分析

因为用于进行大理石花纹提取的牛肉图像主要包括大理石花纹、红肉和背景3部分,所以取聚类数 $C=3$,取模糊加权指数 $m=2$,迭代停止阈值取 $\varepsilon=10^{-3}$ 。

3.1 试验验证

对7幅 640×480 像素且灰度级为 $0 \sim 255$ 的牛肉图像,利用FCM和改进型FCM算法分别进行大理石花纹的提取。对图1所示的牛肉图片提取大理石花纹后的2组图片如图2和图3所示。

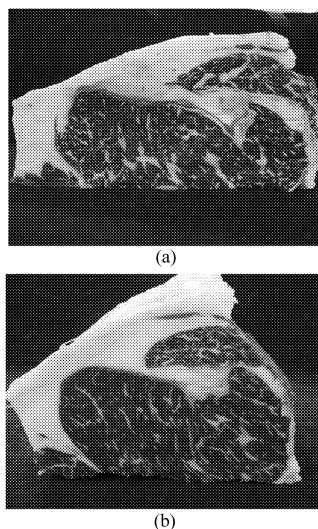


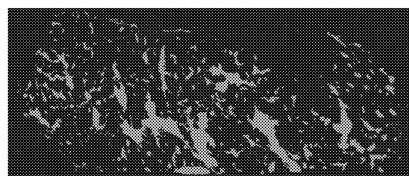
图1 牛肉原图

Fig.1 Original images of beef

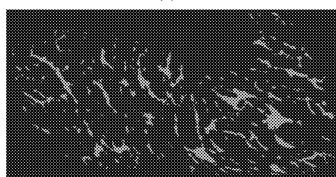
(a) 牛肉1 (b) 牛肉2

3.2 结果分析

为考察算法提取牛肉大理石花纹的准确度,通过牛胴体横切面特征图像数据库中已经标定好等级的标准牛肉图像产生对比参照标准参数,将所有选



(a)

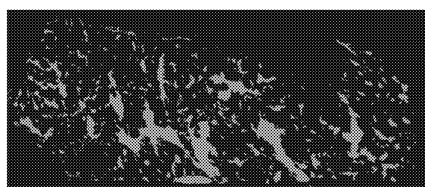


(b)

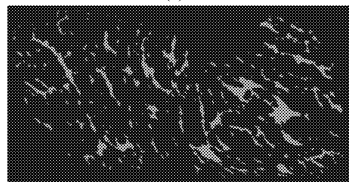
图 2 FCM 算法提取的牛肉大理石花纹

Fig. 2 Beef-marbling extracted by FCM algorithm

(a) 牛肉 1 (b) 牛肉 2



(a)



(b)

图 3 改进型 FCM 算法提取的牛肉大理石花纹

Fig. 3 Beef-marbling extracted by modified FCM algorithm

(a) 牛肉 1 (b) 牛肉 2

定的标准牛肉图像进行相应图像处理, 提取出牛肉大理石花纹的相应参数为对比实际参数。牛肉大理石花纹提取的准确度为

$$a = \frac{T_1}{T_2} \quad (8)$$

式中 T_1 ——算法提取结果与对比参照标准参数重叠的像素数

T_2 ——对比参照标准参数像素数

算法提取结果与对比参照标准参数的像素数越

多, 说明算法的提取准确度越高。表 1 给出了 2 种算法对 7 幅牛肉图像中大理石花纹提取的准确度及平均准确度。

表 1 2 种算法的大理石花纹提取的准确度和平均准确度

Tab.1 Accuracy and average accuracy of the two algorithm

图像编号	FCM 算法	改进型 FCM (FCM-SQP) 算法	%
1	73.8	77.9	
2	74.4	88.1	
3	72.2	84.4	
4	76.7	76.7	
5	79.3	91.4	
6	79.6	92.2	
7	77.4	88.9	
平均准确度	76.2	85.7	

从表 1 可以看出分别运用 FCM 算法和 FCM-SQP 算法进行牛肉大理石花纹提取时, 每一次花纹提取的准确度都在各种算法所对应平均提取准确度的上下小幅度地波动, 说明了改进型模糊 C 均值聚类 (FCM-SQP) 算法在牛肉大理石花纹提取的过程中具有较高的稳定性。运用改进型 FCM (FCM-SQP) 算法所得到的大理石花纹提取平均准确度为 85.7%, 高于传统算法的 76.2%。

4 结束语

根据模糊 C 均值 (FCM) 聚类算法的基本原理研究了一种用于牛肉大理石花纹提取的改进型模糊 C 均值 (FCM-SQP) 聚类算法。通过对传统模糊 C 均值聚类算法中隶属度函数、聚类数 C 和初始聚类中心点选取的优化, 最终实现了改进型模糊 C 均值聚类 (FCM-SQP) 算法, 并通过试验验证了算法。试验结果表明, 在有效地克服了 FCM 算法中聚类中心振荡和数据假聚类等病态现象的基础上, 改进型模糊 C 均值聚类 (FCM-SQP) 算法在牛肉大理石花纹提取的应用中具有很强的鲁棒性和理想的聚类速度, 可以实现牛肉大理石花纹的快速精确提取。

参 考 文 献

- 武军, 刘熙, 邹尔新. 宰后成熟对牛肉肌组织结构及食用品质的影响[J]. 中国兽医学报, 2002, 22(5): 494 ~ 496.
Wu Jun, Liu Xi, Zou Erxin. Effect of maturation after slaughter on the texture of cattle muscle tissue and the beef quality[J]. Chinese Journal of Veterinary Science, 2002, 22(5): 494 ~ 496. (in Chinese)
- 南京农业大学肉类研究室研制. 牛肉等级标准[S]. 2003.
- 贾渊, 姬长英, 汤晓燕. 基于计算机视觉的牛肉分级技术综述[J]. 农业工程学报, 2004, 20(5): 47 ~ 50.
Jia Yuan, Ji Changying, Tang Xiaoyan. Review of beef grading with computer vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2004, 20(5): 47 ~ 50. (in Chinese)
- 陈坤杰, 秦春芳, 姬长英. 牛胴体眼肌切面图像的分割方法[J]. 农业机械学报, 2006, 37(6): 155 ~ 158.
Chen Kunjie, Qin Chunfang, Ji Changying. Segmentation methods used in rib-eye image of beef carcass[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(6): 155 ~ 158. (in Chinese)

- 5 Tadhg Brosnan, Da-Wen Sun. Inspection and grading of agricultural and food products by computer vision systems——a review [J]. Computers and Electronics in Agriculture, 2002, 36(2~3): 193~213.
- 6 Kazuhiko S H, Kenichiro H, Akifumi O, et al. Grading meat quality by image processing [J]. Pattern Recognition, 2000, 33(1): 97~104.
- 7 刘木华, 赵文杰, 张海东. 基于模糊 C 均值聚类的牛肉图像中脂肪和肌肉区域分割技术 [J]. 农业工程学报, 2004, 20(2): 161~163.
Liu Muhua, Zhao Wenjie, Zhang Haidong. Segmentation of fat and lean meat in beef images based on fuzzy C-means clustering [J]. Transactions of the Chinese Society of Agricultural Engineering, 2004, 20(2): 161~163. (in Chinese)
- 8 刘木华, 赵文杰, 张海东. 基于数学形态学的牛肉图像中背长肌分割和大理石花纹提取技术研究 [J]. 农业工程学报, 2004, 20(1): 144~146.
Zhao Wenjie, Liu Muhua, Zhang Haidong. Segmentation of longissimus dorsi and marbling in ribeye imaging based on mathematical morphology [J]. Transactions of the Chinese Society of Agricultural Engineering, 2004, 20(1): 144~146. (in Chinese)
- 9 Chiu Stephen L. A cluster estimation method with extension to fuzzy model identification [C] // Proceedings of the Third IEEE Conference on Control Application, 1994, 2: 1 240~1 245.
- 10 刘宜平, 沈毅, 刘志严. 一种 FCM 聚类算法的改进与优化 [J]. 系统工程与电子技术, 2000, 22(4): 1~3.
Liu Yiping, Shen Yi, Liu Zhiyan. Improvement and optimization of a fuzzy C-means clustering algorithm [J]. Systems Engineering and Electronics, 2000, 22(4): 1~3. (in Chinese)
- 11 李明, 李云松. 改进的快速模糊 C 均值聚类的图像分割方法 [J]. 兰州理工大学学报, 2007, 33(3): 93~99.
Li Ming, Li Yunsong. A modified method for image segmentation with fast fuzzy C-means clustering [J]. Journal of Lanzhou University of Technology, 2007, 33(3): 93~99. (in Chinese)
- 12 陈坤杰, 孙鑫, 陆秋琰. 基于计算机视觉和神经网络的牛肉颜色自动分级 [J]. 农业机械学报, 2009, 40(4): 173~178.
Chen Kunjie, Sun Xin, Lu Qiuyan. Automatic color grading of beef lean tissue based on BP neural network and computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(4): 173~178. (in Chinese)
- 13 伍学千, 廖宜涛, 樊玉霞, 等. 基于 KFCM 和改进分水岭算法的猪肉背最长肌分割技术 [J]. 农业机械学报, 2010, 41(1): 172~176.
Wu Xueqian, Liao Yitao, Fan Yuxia, et al. Segmentation of pork longissimus dorsi based on KFCM clustering and improved watershed algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(1): 172~176. (in Chinese)

~~~~~

(上接第 193 页)

- 7 王永维, 程绍明, 姜雄晖. 温室地下蓄热系统温度的分布试验 [J]. 农业机械学报, 2006, 37(9): 112~115.  
Wang Yongwei, Cheng Shaoming, Jiang Xionghui. Temperature distribution of the underground heat storage system in greenhouse [J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(9): 112~115. (in Chinese)
- 8 王永维, 梁喜凤, 王俊, 等. 温室地下蓄热系统换热管道空气流速对蓄热效果影响 [J]. 农业机械学报, 2009, 40(5): 173~177.  
Wang Yongwei, Liang Xifeng, Wang Jun, et al. Effects of air velocity in heat exchanging pipes on heat saving of the underground heat storage system in greenhouse [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(5): 173~177. (in Chinese)
- 9 马承伟, 苗香雯. 农业生物环境工程 [M]. 北京: 中国农业出版社, 2005.
- 10 周长吉, 杨振声, 冯广和. 现代温室工程 [M]. 北京: 化学工业出版社, 2003.